# Clinicopathologic Features of Non-Small-Cell Lung Cancer with EML4-ALK Fusion Gene.

AUTHOR(S):

Takahashi, Tsuyoshi; Sonobe, Makoto; Kobayashi, Masashi; Yoshizawa, Akihiko; Menju, Toshi; Nakayama, Ei; Mino, Nobuya; ... Okubo, Kenichi; Manabe, Toshiaki; Date, Hiroshi

CITATION:

Takahashi, Tsuyoshi ...[et al]. Clinicopathologic Features of Non-Small-Cell Lung Cancer with EML4-ALK Fusion Gene.. Annals of surgical oncology 2009, 17(3): 889-897

ISSUE DATE:

2009-11-21

URL:

http://hdl.handle.net/2433/108360

# **Title:** Clinicopathologic Features of Non-Small Cell Lung Cancer with the EML4-ALK Fusion Gene

**Running Title:** EML4-ALK fusion gene in resected NSCLC

**Authors:** Tsuyoshi Takahashi,[1] Makoto Sonobe,[1] Masashi Kobayashi,[1] Akihiko Yoshizawa,[2] Toshi Menju,[1] Ei Nakayama,[1] Nobuya Mino,[1] Shotaro Iwakiri,[1] Kiyoshi Sato,[1] Ryo Miyahara,[1] Kenichi Okubo,[1] Toshiaki Manabe,[2] and Hiroshi Date[1]

**Authors' Affiliations:** [1]Department of Thoracic Surgery and [2]Department of Diagnostic Pathology, Faculty of Medicine, Kyoto University, Kyoto, Japan

**Corresponding Author:** Makoto Sonobe, Department of Thoracic Surgery, Faculty of Medicine, Kyoto University. Shogoin-Kawara-cho 54, Sakyo-ku, Kyoto 606-8507, Japan. Phone: +81-75-751-4975, Fax: +81-75-751-4974, E-mail: mysonobe@kuhp.kyoto-u.ac.jp (CC: bridget@kuhp.kyoto-u.ac.jp, msonobe-ths@umin.ac.jp)

**Key Words:** EML4-ALK, non-small cell lung cancer (NSCLC), Epidermal growth factor receptor, KRAS, ERBB2

**[Abstract]**

**Background:** A fusion gene between the echinoderm microtubule-associated protein-like 4 (*EML4*) and the anaplastic lymphoma kinase (*ALK*) has recently been identified in non-small cell lung cancers (NSCLCs). We screened for *EML4-ALK* fusion genes and examined the clinicopathological and genetic characteristics of the fusion-harboring NSCLC tumors.

**Methods:** We examined 313 NSCLC samples from patients who underwent resection at our Hospital between May 2001 and July 2005. We screened for the fusion genes using the reverse transcription polymerase chain reaction (RT-PCR) assay and confirmed the results with direct sequencing. We also examined mutations in the epidermal growth factor receptor *(EGFR)*, *KRAS*, and *ERBB2* genes.

**Results:** Five *EML4-ALK* fusion genes were detected (four from 111 female samples and one from 202 male samples; 1.6% overall). All five genes were found in adenocarcinomas and accounted for 2.4% of the 211 adenocarcinoma samples. One *EML4-ALK* fusion was variant 1, and two were variant 3. In addition, we also found two new fusion variants. Patients with fusion-positive tumors were non-smokers or light smokers. Among the 211 adenocarcinomas, mutations in *EGFR*, *KRAS*, and *ERBB2* were detected in 105, 29, and 7 tumors, respectively. Interestingly, all of the fusion-positive NSCLCs had no mutations within these genes.

**Conclusions:** The *EML4-ALK* fusion genes were observed predominantly in adenocarcinomas, in female or non-smoking populations. Additionally, the *EML4-ALK* fusions were mutually exclusive with mutations in the *EGFR*, *KRAS*, and *ERBB2* genes.

## [Introduction]

Primary lung cancer is the leading cause of cancer-related death in many industrialized countries. Although a combination of surgery, chemotherapy, and/or radiotherapy can be used to treat this disease, the prognosis for patients remains dismal. In an effort to develop new treatment strategies, several molecular-targeting drugs have recently been developed. It is known that epidermal growth factor receptor (EGFR*)* gene mutations are one of the causes of primary lung adenocarcinomas, and to some extent, tyrosine kinase inhibitors that target EGFR (*e.g.*, gefitinib and erlotinib) have proven successful for controlling this disease.[1-11] Therefore, the identification of the key oncogenes for lung cancer is a very important step toward the development of novel molecular-targeting agents. Interestingly, *EGFR* gene mutations have been revealed to be exclusive in adenocarcinomas, and most of these patients are from the Asian, female, non-smoker population.[6-9, 11-22, 24] Gefitinib is most commonly applied to lung adenocarcinomas with these typical clinicopathological features.

These facts underscore the importance of discovering favorable clinicopathologic and genetic characteristics for each type of tyrosine kinase inhibitor therapy.

Recently, Soda and colleagues (2007) identified another type of tyrosine kinase with accelerated activity in a fusion gene between the echinoderm microtubule-associated protein-like 4 (*EML4*) and the anaplastic lymphoma kinase (*ALK*).[23] Indeed, ALK kinase inhibitors have already been developed and have been reported to suppress the growth of *EML4-ALK* fusion positive cells.[23, 24, 26, 27] Thus, treatment with ALK inhibitors can be effective for non-small cell lung cancer (NSCLC) patients whose tumors contain an *EML4-ALK* fusion.

In this report, we examine the *EML4-ALK* fusion gene in a relatively large number of clinical samples, and investigate its clinicopathologic and genetic background, in order to identify useful information on patient selection for ALK inhibitor therapy.

## [Materials and Methods]

### *Patients and sample collection*

This study included 313 Japanese NSCLC patients (211 adenocarcinomas, 75 squamous cell carcinomas, 17 large cell carcinomas, and 10 other NSCLCs) who underwent

pulmonary resection at Kyoto University Hospital between May 2001 and July 2005. The pathological stage was evaluated according to the current international TNM staging system and the classification of the World Health Organization (WHO). Patient data were obtained from inpatient and outpatient medical records. All patients were followed-up (either to the time of death or through April 2009) using outpatient medical records and by letter or phone. The genetic analyses in the present study were approved by the Ethics Committee of our institute.

***Preparation of cDNA***

All of the tumor tissues were stored in RNAlater TissueProtect Tubes (Qiagen) immediately after pulmonary resections for RNA stabilization. Total RNA was extracted from specimens using the RNeasy Mini Kit (Qiagen), and the RNA extract was incubated with RNase-free DNase I (Qiagen) to remove contaminating DNA. Reverse-transcription of total RNA was performed using Ready-To-Go You-Prime First-Strand Beads (Amersham Bioscience) to generate cDNA.

***Detection of EML4-ALK fusion genes***

For detection of *EML4-ALK* fusion cDNAs, we performed reverse transcription polymerase chain reaction (RT-PCR) assays using the primer set described by Soda *et al.*:[23, 27]

5’-GTGCAGTGTTTAGCATTCTTGGGG-3’ (forward primer, on exon 13 of *EML4*) and

5’-TCTTGCCAGCAAAGCAGTAGTTGG-3’ (reverse primer, on exon 21 of *ALK*). We also used the following primer set to detect other types of fusion transcripts consisting of the upper exons of *EML4* and *ALK*: 5’-GTCAGCTCTTGAGTCACGAGTT-3’ (forward primer, on exon 2 of *EML4*) and 5’-TCTTGCCAGCAAAGCAGTAGTTGG-3’ (reverse primer, on exon 21 of *ALK*). Amplification was initiated by pre-incubation for 15 min at 95°C for the initial activation, followed by 35 cycles of: denaturation for 30 s at 95°C, primer annealing for 30 s at 66°C, and elongation for 1 min at 72°C using HotStar *Taq* Master Mix Kit (Qiagen). A 12μL aliquot of each reaction mixture was electrophoresed through a 2% agarose gel and stained by ethidium bromide for visualization under ultraviolet light. We used the primers 5’-ACAACAGCCTCAAGATCATCAG-3’ (forward) and 5’-TCTTCTGGGTGGCAGTGATG-3’ (reverse) to amplify the *GAPDH* gene as an internal control. *GAPDH* amplification was performed by pre-incubation for 10 min at 95°C for initial activation, followed by 35 cycles of: denaturation for 15 s at 94°C, primer annealing for 15 s at 59°C, and elongation for 15 s at 72°C.

***Detection of the mutations within the EGFR, KRAS, and ERBB2 genes***

As reported previously, we analyzed mutations in the *EGFR* (exons 18-21)[28] and *ERBB2* (exons 19-20) genes [30] by the PCR-single strand conformational polymorphism (PCR-SSCP) technique. We also used a modified mutagenic restriction fragment length

polymorphism (PCR-RFLP) method to screen for mutations in codon 12 of the *KRAS* gene.[28]

***Histopathological analysis***

All of the surgical specimens were subjected to hematoxylin-eosin staining at the Department of Diagnostic Pathology of our hospital. Two board-certified pathologists independently reviewed the slides and performed diagnoses according to the 2004 WHO Classification of lung tumors.

**[Results]**

***Detection of the EML4-ALK fusion gene and analysis of fusion patterns***

Using RT-PCR, five fusion transcripts were detected (Fig. 1). The control *GAPDH* gene was amplified in all of the 313 samples (data not shown). The *EML4* (exon 2)-*ALK* primer set amplified *EML4-ALK* fusion genes from all five of these patients (1-5), and the *EML4* (exon 13)-*ALK* primer set amplified fusion genes from patients 1, 4, and 5. Direct sequence analysis confirmed the variant that was harbored by each tumor (Table 1). As a control, no gene was amplified by both primer sets from A549 cells, which do not carry the *EML4-ALK* fusion gene. Tumors from patient 4 carried a new fusion variant, which involved the fusion of full length exon 20 of *EML4* and full length exon 20 of *ALK* with an 18-bp insertion sequence (5'-CTGACCACCCACCTGCAG-3') between the two exons. This inserted sequence is identical to the last 18-bp of *ALK* intron 19, which means that the break point of ALK was located at an upstream intron (Fig. 2A, B).

The tumor from patient 5 also had a novel transcript that involved a fusion between full-length *EML4* exon 17 and full-length *ALK* exon 20, with a 68-bp insertion sequence (5'-AGTCTTGCTCTGTCTCCCAGGCTGGAGTGCAGTGGCAATTTACACATTTCAATTCATTCGATCCTCAG-3') (Fig. 2A, C). Six sequences identical to the first part of the 14-bp sequence (5'-AGTCTTGCTCTGTC-3') were found in *ALK* introns: one sequence in intron 1, three in intron 3, one in intron 4, and one in intron 8. The reversed sequence (5'-GACAGAGCAAGACT-3') was also found in *ALK* introns: one each in introns 1, 3, 4, 5, and 26. Thus 11 copies of this 14-bp sequence were scattered throughout the *ALK* introns. One identical sequence to the middle part of the inserted sequence (5'-TCCCAGGCTGGAGTGCAGT-3', 19-bp) was also found in *ALK* intron 4, and its reversed sequence (5'-ACTGCACTCCAGCCTGGGA-3') was found in *ALK* intron 5. Interestingly, we found nineteen sequences resembling this middle inserted sequence in *ALK* introns, but they lacked the first thymidine and the last three nucleotides (*i.e.*, 5'-CCCAGGCTGGAGTGC-3'). We also found twenty-one reversed copies of this sequence in *ALK* introns, thus *ALK* introns had 42 copies of 5'-CCCAGGCTGGAGTGC-3' and its reversed sequence. Among them, thirteen copies of 5'-CCCAGGCTGGAGTGCAGT-3' and ten of the reversed sequence 5'-ACTGCACTCCAGCCTGGG-3' were found. In *ALK* intron 19, one identical sequence to the last part of the inserted sequence (5'-GGCAATTTACACATTTCAATTCATTCGATCCTCAG-3', 35-bp) was also observed, and its reversed sequence (5'-CTGAGGATCGAATGAATTGAAATGTGTAAATTGCC-3') was found in *ALK* intron 19. Thus, the inserted sequence of patient 5 consists of three ALK intron sequences. The *EML4-ALK* fusions from patients 4 and 5 resemble variant 3b, in which exon 6 of *EML* is fused to exon 20 of *ALK* with a 33-bp insertion.

***Patients with EML4-ALK fusion-positive NSCLCs***

Out of 313 patients, 5 (1.6%) had tumors that contained *EML4-ALK* fusion genes. These five fusion-positive tumors were adenocarcinomas, and the clinicopathologic profiles of the five

fusion-positive *vs.* negative adenocarcinoma patients are shown in Table 2. No patients with *EML4-ALK* fusion-positive tumors had received any induction chemotherapy, because their clinical stages had been evaluated as 1A to 2A. Two fusion-positive adenocarcinomas were detected from pathological stage 1A tumors and three from stage 3A tumors. Four of the five patients with fusion-positive tumors were nonsmokers, and the only smoker had a 0.25 pack/year smoking history, which was less than fusion-negative patients that smoked.

***Mutations of the EGFR and KRAS genes in NSCLCs***

Out of 313 samples, *EGFR* mutations were detected in 105 adenocarcinomas and in one squamous cell carcinoma. *KRAS* mutations were detected in 35 patients: 29 in adenocarcinomas, three in large cell carcinomas, two in squamous cell carcinomas, and one in pleomorphic carcinoma. *ERBB2* mutations were detected in seven adenocarcinomas and in one squamous cell carcinoma (Table 3). All five fusion-positive adenocarcinomas had wild-type *EGFR*, *KRAS*, and *ERBB2* genes (Table 2). Further, tumors with the *EML4-ALK* fusion had no *p53* (exons 5-8) gene mutations (data not shown).

***Prognosis and overall survival of the EML4-ALK fusion-harboring patients***

Two patients with p-stage 1A adenocarcinoma received no adjuvant therapy and survived for over 66 months (patient 1) and 60 months (patient 2) without any recurrence or metastasis (Table 1). A patient with p-stage 3A adenocarcinoma (patient 3) underwent a

right upper lobectomy, followed by two cycles of combined adjuvant chemotherapy with a carboplatin plus paclitaxel regimen. A metastatic lesion appeared in the remaining middle lobe, and a right middle lobectomy was performed 73 months after the primary lung resection. Histopathological studies determined that the tumor in the resected right middle lobe and the anterior mediastinal lymph node (#3a) were a recurrent tumor and metastatic lymph node of the primary tumor. Additional RT-PCR assays revealed that the relapsed tumor cells in the right middle lobe harbored the same *EML4-ALK* fusion. Histopathological studies demonstrated that this second tumor had the same characteristics as the primary resected tumor. With 100 mg/day of S-1 (Tegafur/ Gimeracil/ Oteracil) medication, patient 3 survived for over 99 months following the first pulmonary resection (26 months after the second resection).

Two patients with p-stage 3A adenocarcinoma (patients 4 and 5) underwent a left upper lobectomy and a left lower lobectomy, respectively. Patient 4 underwent two cycles of adjuvant chemotherapy with the carboplatin plus paclitaxel regimen. Multiple brain metastases appeared 16 months after the pulmonary resection, and whole brain radiation therapy (30 Gy/ 10 fractures) was performed. Patient 4 died 28 months after the operation. Eight months after the lobectomy, chest CT studies revealed that patient 5 had a metastatic nodule in the right lower lung and metastatic lymph nodes in the left mediastinum. This

patient underwent two cycles of chemotherapy with the carboplatin plus paclitaxel regimen. Seventeen months after the resection, a metastatic lesion was found in the brain. Patient 5 then underwent gamma knife radio-surgery with vinorelbine. Metastatic lesions in the mediastinum and liver grew, and patient 5 died 53 months after the pulmonary resection.

***Histopathological analysis of EML4-ALK fusion-harboring tumors***

One of the five fusion-harboring adenocarcinomas was an acinar adenocarcinoma, and the other four were classified as mixed subtype adenocarcinomas: two of the acinar predominant subtype and two of the papillary predominant subtype. The four adenocarcinomas of mixed subtype were evaluated as poorly differentiated (Grade 3), and the acinar adenocarcinoma was graded as moderately differentiated (Grade 2) (Fig. 3).

**[Discussion]**

Using RT-PCR analysis, we detected five *EML4-ALK* fusion genes from among 313 NSCLC specimens (1.6%). The frequency rates for *EML4-ALK* fusions in NSCLCs have been reported in several studies,[23, 27, 31-35] and our results are comparable with the previously published data. In our study, samples were collected entirely from a Japanese population; the frequency of *EML4-ALK* fusions was similar to that of U.S. plus Korean,[27] European,[36] or

Chinese[37] populations. This finding suggests that there is no significant difference in the frequency of *EML4-ALK* fusion genes between countries.

The prognosis and the clinical course after primary tumor resection of our patients with *EML4-ALK* fusion-harboring NSCLC seemed to be relatively favorable. Both of the patients with pathological grade 1A adenocarcinomas survived for over 60 months without any recurrence or metastasis. One with a pathological grade 3A adenocarcinoma (Patient 2) survived for over 99 months, with pulmonary metastasis occurring 73 months after the primary lung resection. This case showed that an *EML4-ALK* fusion positive tumor could relapse after a long interval (>60 months), but adequate local and systemic control could provide an excellent survival benefit. One patient with a pathological grade 3A adenocarcinoma had brain metastasis 8 months after the operation, yet still survived 53 months. Inamura and colleagues (2007) reported six patients with pathological grade 1A adenocarcinomas and two patients with grade 3A adenocarcinomas who survived over 59.0 ± 8.8 months and 39 ± 6.9 months, respectively.[35] Because there are not enough clinical data dealing with fusion-harboring patients thus far, it is difficult to appreciate the prognosis of *EML4-ALK* fusion-positive patients statistically. However, considering the Inamura and colleagues (2007) report and our data, genetic alterations producing *EML4-ALK* fusions may not be an unfavorable prognostic factor for adenocarcinoma patients.

We investigated the incidence of *EML4-ALK* in this relatively large-scale screen and analyzed its characteristics according to the age, sex, and smoking habits of the patients, as well as the histology, differentiation, and p-stage of the tumors. Of the five *EML4-ALK* fusion-harboring adenocarcinoma patients, four were female. The fusion-positive patients were mostly non-smokers; only one patient had a smoking history of a 0.25 pack/year. It should be noted that the pack/year number of fusion-positive patients was smaller than that of fusion-negative patients (Table 2; 0.05 ± 0.01 *vs.* 29.6 ± 36.6, $p<0.001$).

All five fusions were detected only in adenocarcinomas (Table 3). These findings may indicate that the *EML4-ALK* fusion gene is related to the oncogenes that have been mutated in some nonsmokers' lung adenocarcinomas. This tendency is compatible with several previously reported results.[25, 35, 37] This finding also suggests the importance of patient selection for gene-targeted therapy using ALK kinase inhibitors, because the efficiency of tyrosine kinase inhibition is strongly related to typical clinicopathologic backgrounds, as is recognized for the use of gefitinib and erlotinib. Our data and that of previous investigators indicate the specific clinicopathologic features of candidates for ALK tyrosine kinase inhibitor therapy.

As shown in this and other studies,[10-12, 24, 28, 38-45] about half of Japanese lung adenocarcinomas harbor a mutation in *EGFR*, 10-15% in *KRAS*, and 4-6% in *ERBB2*. Our

genetic analysis of the *EGFR*, *KRAS* and *ERBB2* genes demonstrated that the *EML4-ALK* fusion gene is mutually exclusive with mutations in these genes. Therefore, two-thirds of lung adenocarcinomas are classified based on the genetic alterations of their tyrosine kinase, which indicates that most adenocarcinomas are candidates for gene-targeted therapy, especially therapy involving tyrosine kinase inhibitors.

Usually, gefitinib or erlotinib therapy has been applied to control locally advanced or metastatic adenocarcinomas after several regimens of chemotherapy. Up to the present, most of adenocarcinoma treated with these tyrosine kinase inhibitors have not been confirmed their *EGFR* gene status because screening for mutations in the *EGFR* has not been practical due to financial problems and/or technical inconvenience. According to the findings from many researchers and our study, good responders to gefitinib or erlotinib can be estimated to harbor mutations in the *EGFR* gene and not the *EML4-ALK* fusions. At present, as gefitinib or erlotinib are usually applied for certain adenocarcinoma patient groups with specific clinical characteristics (e.g., Asian, female, and non-smokers), their efficiency are getting more favorable. Nevertheless, some non-responders would be left and a certain number of *EML4-ALK* fusion-harboring adenocarcinomas would hide behind them. Therefore, we considered that it would be very beneficial to demonstrate RT-PCR assays for detection of the *EML4-ALK* fusions in tumors of the non-responders to such conventional

*EGFR* kinase inhibitors. We believe this easy and economical screening method and the future ALK kinase targeted therapy would be very significant for treatment of the *EML4-ALK* fusion-harboring adenocarcinomas.

Nine *EML4-ALK* variants have been previously identified by other investigators: 1, 2, 3a, 3b, 4, 5a, 5b, 6, and 7. Variants 1, 2, 3a, and 5a consist of a simple fusion between an exon of *EML4* (exons 13, 20, 6, and 2, respectively) and exon 20 of the *ALK* gene. Variant 3b involves a 33-bp insertion between exon 6 of *EML4* and exon 20 of *ALK*, and variant 5b involves a 117-bp insertion between exon 2 of *EML4* and exon 20 of *ALK*. In variant 4, the *EML4* exon 14 is fused via an 11-bp insertion to the 50th nucleotide of *ALK* exon 20. *EML4* exon 13 is fused 69-bp upstream of *ALK* exon 20 in variant 6, and *EML4* exon 14 is fused to nucleotide 13 of *ALK* exon 20 in variant 7.[46] Here, we have identified two novel *EML4-ALK* fusion variants. In the tumor of patient 4, *EML4* exon 20 was fused via an 18-bp sequence to *ALK* exon 20. The inserted sequence of this variant comprises the last 18-bp of *ALK* intron 19. Thus, this fusion can be classified as a subtype of variant 2. Compared with variant 2, the fusion sequence of patient 5 had a 275-bp deletion of full length of *EML4* exon 18 (89-bp), exon 19 (98-bp), and exon 20 (88-bp), and a 68-bp insertion. Therefore, the fusion from patient 5 had a 207-bp smaller sequence than variant 2, which does not provide frame shift. This type of fusion was unique, because it contained mixed sequences from *ALK*

introns between the *EML4* and *ALK* exons. The characteristic phenomenon of this variant was that the inserted sequences were composed of three *ALK* intron sequences. We could not find *ALK* intron components in the other types of variants containing inserted sequence between *EML4* and *ALK*. Above all, the most peculiar features of the fusion gene from patient 5 were eleven sequences identical to the first 14-bp, and another 42 of 15-bp, *ALK* intron sequences and their reversed sequences. Reversed sequences were correlated with the existence of identical sequences on anti-sense DNA, scattered around the *ALK* introns. There is no confirmed evidence, but this uncommon event of quite a number of duplicated sequences may induce altered *EML4-ALK* fusion conformation. Unfortunately, due to the small number of *EML4-ALK* fusion samples, the influence of fusion variations on clinicopathologic features remains unknown.

According to the WHO classification of lung tumors (2004), adenocarcinomas can be classified as acinar and papillary adenocarcinomas, bronchioloalveolar carcinoma, solid adenocarcinoma with mucin production, mixed subtype adenocarcinomas, and their variants. In our study, one fusion-positive sample was an acinar carcinoma, and four were adenocarcinomas of mixed subtype (two of the acinar predominant type and two of the papillary predominant type). Inamura and colleagues (2009) reported the detection of five fusion-positive adenocarcinomas out of 206 papillary predominant subtypes and six out of

34 acinar predominant subtypes,[35] showing a significantly higher tendency for fusions in acinar-predominant subtype adenocarcinomas. Considering these findings, *EML4-ALK* fusion-harboring adenocarcinomas are likely to be related to acinar components.

In conclusion, we determined that *EML4-ALK* fusion-harboring lung cancers were only present in adenocarcinomas and in patients with little or no smoking history in our test group. In addition, we concluded that the *EML4-ALK* fusions were mutually exclusive with mutations in the *EGFR*, *KRAS,* and *ERBB2* genes. These observations provide additional information on favorable responders with respect to specific ALK kinase inhibitor therapy.

**[Acknowledgements]**

We thank Ms. Seiko Sakai for excellent secretarial assistance. This project was not supported by any outside funds.

**[References]**

**1.** Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004; 304(5676):1497-500.

**2.** Lee DH, Han JY, Yu SY, et al. The role of gefitinib treatment for Korean never-smokers with advanced or metastatic adenocarcinoma of the lung: a prospective study. J Thorac Oncol 2006; 1(9):965-71.

**3.** Tamura K, Okamoto I, Kashii T, et al. Multicentre prospective phase II trial of gefitinib for advanced non-small cell lung cancer with epidermal growth factor receptor mutations: results of the West Japan Thoracic Oncology Group trial (WJTOG0403). Br J Cancer 2008; 98(5):907-14.

**4.** Qin BM, Chen X, Zhu JD, Pei DQ. Identification of EGFR kinase domain mutations among lung cancer patients in China: implication for targeted cancer therapy. Cell Res 2005; 15(3):212-7.

**5.** Pao W, Miller V, Zakowski M, et al. EGF receptor mutations are common in lung cancers from “never-smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci USA 2004; 101(36):13306-11.

**6.** Fukui T, Mitsudomi T. Mutations in the epidermal growth factor receptor gene and effects of EGFR-tyrosine kinase inhibitors on lung cancers. Gen Thorac Cardiovasc Surg 2008; 56(3):97-103.

**7.** Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004; 304(5676):1497-500.

**8.** Mitudomi T, Yatabe Y. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. Cancer Sci 2007; 98(12):1817-24.

**9.** Marchetti A, Martella C, Felicioni L, et al. EGFR mutations in non-small-cell lung cancer: analysis of a large series of cases and development of a rapid and sensitive method for diagnostic screening with potential implications on pharmacologic treatment. J Clin Oncol 2005; 23(4):857-65.

**10.** Huang SF, Liu HP, Li LH, et al. High frequency of epidermal growth factor receptor mutations with complex patterns in non-small cell lung cancers related to gefitinib responsiveness in Taiwan. Clin Cancer Res 2004; 10(24):8195-203.

**11.** Taron M, Ichinose Y, Rosell R, et al. Activating mutations in the tyrosine kinase domain of the epidermal growth factor receptor are associated with improved survival in gefitinib-treated chemorefractory lung adenocarcinomas. Clin Cancer Res 2005; 11(16):5876-85.

**12.** Mu XL, Li LY, Zhang XT, et al. Gefitinib-sensitive mutations of the epidermal growth factor receptor tyrosine kinase domain in Chinese patients with non-small cell lung cancer. Clin Cancer Res 2005; 11(12):4289-94.

**13.** Lee DH, Han JY, Yu SY, et al. The role of gefitinib treatment for Korean never-smokers with advanced or metastatic adenocarcinoma of the lung: a prospective study. J Thorac Oncol 2006; 1(9):965-71.

**14.** Wu YL,Zhong WZ, Li LY, et al. Epidermal growth factor receptor mutations and their correlation with gefitinib therapy in patients with non-small cell lung cancer: a meta-analysis based on updated individual patient data from six medical centers in mainland of China. J Thorac Oncol 2007; 2(5):430-9.

**15.** Chang CH, Chen KY, Young-Xu Y, et al. The safety and efficiency of gefitinib versus platinum-based doublets chemotherapy as the first-line treatment for advanced non-small-cell lung cancer patients in East Asia: a meta-analysis. Lung Cancer 2008; 62(2):242-52.

**16.** Haringhuizen A, van Tinteren H, Vaessen HF, Bass P, van Zandwijk N. Gefitinib as a last treatment option for non-small-cell lung cancer: durable disease control in a subset of patients. Ann Oncol 2004; 15(5):786-92.

**17.** Byrne BJ, Garst J. Epidermal growth factor receptor inhibitions and their role in non-small-cell lung cancer. Curr

Oncol Rep 2005; 7(4):241-7.

**18.** Tamura K, Fukuoka M. Gefitinib in non-small cell lung cancer. Expert Opin Pharmacother 2005; 6(6):985-93.

**19.** Park K, Goto K. A review of the benefit-risk profile of gefitinib in Asian patients with advanced non-small-cell lung cancer. Curr Med Res Opin 2006; 22(3):561-73.

**20.** Lin WC, Chiu CH, Liou JL, Chen YM, Perng RP, Tsai CM. Gefitinib as front-line treatment in Chinese patients with advanced non-small-cell lung cancer. Lung Cancer 2006; 54(2):193-9.

**21.** Tsao AS, Tang XM, Sabloff B, et al. Clinicopathologic characteristics of the EGFR mutations in non-small cell lung cancer. J Thorac Oncol 2006; 1(3):231-9.

**22.** Kosaka T, Yatabe Y, Endoh H, Kuwano H, Takahashi T, Mitsudomi T. Mutations of the epidermal growth factor gene in lung cancer: biological and clinical implications. Cancer Res 2004; 64(24):8919-23.

**23.** Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007; 448(7153):561-6.

**24.** Mitsudomi T, Kosaka T, Endoh H, et al. Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. J Clin Oncol 2005; 23(11):2513-20.

**25.** Soda M, Takada S, Takeuchi K, et al. A mouse model for EML4-ALK-positive lung cancer. Proc Natl Acad Sci USA 2008; 105(50):19893-7.

**26.** Choi YL, Takeuchi K, Soda M, et al. Identification of novel isoforms of the EML4-ALK transforming gene in non-small cell lung cancer. Cancer Res 2008; 68(13):4971-6.

**27.** Koivunen JP, Mermel C, Zejnullahu K, et al. EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer. Clin Cancer Res 2008; 14(13):4275-83.

**28.** Takeuchi K, Choi YL, Soda M, et al. Multiplex reverse transcription-PCR screening for EML4-ALK fusion transcripts. Clin Cancer Res 2008; 14(20):6618-24.

**29.** Sonobe M, Manabe T, Wada H, Tanaka F. Mutations in the epidermal growth factor receptor gene are linked to smoking-independent, lung adenocarcinoma. Brit J Cancer 2005; 93(3):355-363.

**30.** Sonobe M, Manabe T, Wada H, Tanaka F. Lung Adenocarcinoma Harboring Mutations in the ERBB2 Kinase Domain. JMD 2006; 8:351-356.

**31.** Fukuyoshi Y, Inoue H, Kita Y, et al. EML4-ALK fusion transcript is not found in gastrointestinal and breast cancers. Br J Cancer 2008; 98(9):1536-9.

32. Perner S, Wagner PL, Demichelis F, et al. EML4-ALK fusion lung cancer: a rare acquired event. Neoplasia 2008; 10(3):298-302.

**33.** Shinmura K, Kageyama S, Tao H, et al. EML4-ALK fusion transcripts, but no NPM-, TPM3-, CLTC-, ATIC-, or TFG-ALK fusion transcripts, in non-small cell lung carcinomas. Lung Cancer 2008; 61(2):163-9.

**34.** Inamura K, Takeuchi K, Togashi Y, et al. EML4-ALK fusion is linked to histological characteristics in a subset of lung cancers. J Thorac Oncol 2008; 3(1):13-7.

**35.** Inamura K, Takeuchi K, Togashi Y, et al. EML4-ALK lung cancers are characterized by rare other mutations, a TTF-1 cell lineage, an acinar histology, and young onset. Mod Pathol 2009; 22(4):508-15.

**36.** Martelli MP, Sozzi G, Hernandez L, et al. EML4-ALK rearrangement in non-small cell lung cancer and non-tumor lung tissues. Am J Pathol 2009; 174(2):661-70.

**37.** Wong DW, Leung EL, So KK, et al. The EML4-ALK fusion gene is involved in various histologic types of lung cancers from nonsmokers with wild-type EGFR and KRAS. Cancer 2009; 115(8):1723-33.

**38.** Sasaki H, Shimizu S, Endo K, et al. EGFR and erbB2 mutation status in Japanese lung cancer patients. Int J Cancer 2006; 118(1):180-4.

**39.** Takano T, Fukui T, Ohe Y, et al. EGFR mutations predict survival benefit from gefitinib in patients with advanced lung adenocarcinoma: a historical comparison of patients treated before and after gefitinib approval in Japan. J Clin Oncol 2008; 26(34):5589-95.

**40.** Stephens P, Hunter C, Bignell G, et al. Lung cancer: intragenic ERBB2 kinase mutations in tumors. Nature 2004; 431(7008):525-6.

**41.** Shigematsu H, Takahashi T, Nomura M, et al. Somatic mutations of the HER2 kinase domain in lung adenocarcinomas. Cancer Res 2005; 65(1):1642-6.

**42.** Sasaki H, Shimizu S, Endo K, et al. EGFR and erbB2 mutations status in Japanese lung cancer. Int J Cancer 2006; 118(1):180-4.

**43.** Noda N, Matsuzoe D, Konno T, Kawahara K, Yamashita Y, Shirakusa T. K-ras gene mutations in non-small cell lung

cancer in Japanese. Oncol Rep 2001; 8(4):889-92.

**44.** Shibata T, Hanada S, Kokubo A, et al. Gene expression profiling of epidermal growth factor receptor/KRAS pathway activation in lung adenocarcinoma. Cancer Sci 2007; 98(7):985-91.

**45.** Sonobe M, Nakagawa M, Takenaka K, et al. Influence of epidermal growth factor receptor (EGFR) gene mutations on the expression of EGFR, phosphoryl-Akt, and phosphoryl-MAPK, and on the prognosis of patients with non-small cell lung cancer. J Surg Oncol 2007; 95:63-69.

**46.** Takeuchi K, Choi YL, Togashi Y, et al. KIF5B-ALK, a novel fusion oncokinase identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer. Clin Cancer Res 2009; 15(9):3143-9.

## [Figure legends]

### Figure 1A

RT-PCR studies using the *EML4* (exon 2)-*ALK* primer set. The λ DNA-*Hin*dIII Digest (BioLabs) and the φX174 *Hae*III digest (Takara Bio) were used as the DNA ladder markers. In each patient (1-5), a single band was amplified, and its size was subsequently determined: 1357-2027 bp band in patient 1, 603-872 bp band in patient 2, 603-872 bp band in patient 3, 2027-2322 bp band in patient 4, and 2027-2322 bp band in patient 5.

### Figure 1B

RT-PCR studies using the *EML4* (exon13)-*ALK* primer set. The same DNA ladder markers as above were used. A single band was amplified in three of the five patients, and the sizes were subsequently determined: a <281 bp band in patient 1, 872-1078 bp band in patient 4, and 603-872bp band in patient 5. No RT-PCR products were obtained in patients 2 and 3.

**Figure 2A**

Schema showing two novel *EML4-ALK* fusion variants and variant 2. In patient 4, exon 20 of *EML4* is fused via an 18-bp sequence to exon 20 of *ALK*. This fusion could be classified as a subtype of variant 2. In patient 5, exon 17 of *EML4* is fused via a 68-bp insertion to exon 20 of *ALK*. Compared with variant 2, this fusion type had a full length deletion *EML4* exon 18-20 (275-bp) and a 68-bp insertion. This results in a 207-bp shorter fusion sequence than variant 2, which provided no frame shift.

**Figure 2B**

Schema showing the fusion pattern of patient 4 with precise sequence information. The inserted sequence was identical to the last part of ALK intron 19, which indicates that the break point of *ALK* is 18-bp upstream of that in variant 2.

**Figure 2C**

Schema showing the location of identical sequences composing inserted sequences. Three parts of the inserted sequence are named A, B, and C. Sequences marked as “A” and a left

arrow indicate the sequence 5'-AGTCTTGCTCTGTC-3', and those marked as "A" and a right arrow indicate the reversed sequence 5'-GACAGAGCAAGACT-3'. Sequences marked as "B" with a left arrow indicate the sequence 5'-TCCCAGGCTGGAGTGCAGT-3', and those marked by "B" with a right arrow indicate the reversed sequence. Sequence marked as "C" with a left arrow indicates the 35-bp sequence 5'-GGCAATTTACACATTTCAATTCATTCGATCCTCAG-3'.

**Figure 3**

Histopathological studies of *EML4-ALK* fusion-harboring tumors. All of the carcinomas were adenocarcinomas. One was an acinar adenocarcinoma (patient 2), and the other four were classified as adenocarcinomas of mixed subtypes. In patients 3 and 4, acinar characteristics were evaluated to be predominant. The acinar adenocarcinoma in patient 2 was classified as moderately differentiated, and the others were all classified as poorly differentiated (grade 3) types.

**Table 1. Clinical Features and Details of EML4-ALK Fusion Gene Positive Patients**

| Pt | Age | Sex | Histology and Subtype of Adenocarcinomas | Predominant subtype | diff. | p-stage | Tumor location | EML4-ALK variants | Smoking (pack-year) | Overall survival (M) | Prognosis | Disease-free survival (M) | Site of recurrence | EGFR mutation | KRAS mutation | ERBB2 mutation |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 77 | F | Ad with mixed subtypes | Papillary | Poor | 1A | LUL | V1 | 0 | 66M | alive | 66M | none | - | - | - |
| 2 | 72 | F | Acinar Ad | Acinar | Mod | 1A | LUL | V3b | 0 | 60M | alive | 60M | none | - | - | - |
| 3 | 53 | M | Ad with mixed subtypes | Acinar | Poor | 3A | RUL | V3a | 0.25 | 99M | alive | 73M | lung | - | - | - |
| 4 | 75 | F | Ad with mixed subtypes | Acinar | Poor | 3A | LUL | novel type | 0 | 28M | dead | 16M | brain | - | - | - |
| 5 | 73 | F | Ad with mixed subtypes | Papillary | Poor | 3A | LLL | novel type | 0 | 53M | dead | 8M | brain | - | - | - |

Pt, patient; Ad, adenocarcinoma; diff, differentiation; Mod, moderately differentiated; p-stage, pathological stage; LUL, left upper lobe; RUL, right upper lobe; RLL, right lower lobe; LLL, left lower lobe; V, EML4-ALK variant; M, months; -, not detected

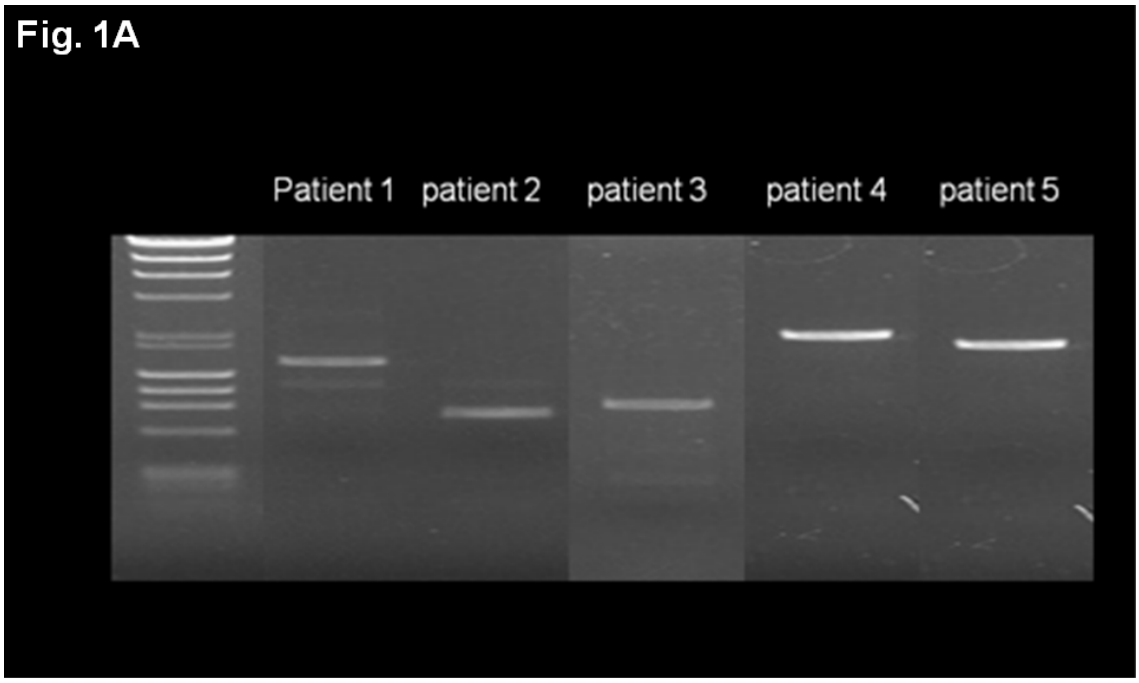
**Table 2. Clinicopathologic and Genetic Comparisons Between EML4-ALK Fusion-positive and -negative Adenocarcinomas**

| Characteristic | | Total | EML4-ALK fusion (+) (n=5) | EML4-ALK fusion (-) (n=206) | *P* |
|---|---|---|---|---|---|
| Age (yr; mean) | | | 70.0 ± 9.7 | 65.2 ± 10.1 | 0.29 |
| Sex | | | | | |
| | Male | 111 | 1 | 110 | 0.31 |
| | Female | 100 | 4 | 96 | |
| Smoking | | | | | |
| | Never | 92 | 4 | 88 | 0.58 |
| | Smoker | 119 | 1 | 118 | |
| | Pack-year (mean ± SD) | | 0.05 ± 0.11 | 29.6 ± 36.6 | p<0.001 |
| Differentiation | | | | | |
| | Well | 87 | 0 | 86 | p<0.01† |
| | Moderate | 91 | 1 | 87 | |
| | Poor | 33 | 4 | 33 | |
| p-Stage | | | | | |
| | I | 141 | 2 | 139 | 0.16‡ |
| | II | 20 | 0 | 20 | |
| | III | 42 | 3 | 39 | |
| | IV | 8 | 0 | 8 | |
| EGFR | | | | | |
| | Wild type | 106 | 5 | 101 | 0.072 |
| | Mutated | 105 | 0 | 105 | |
| KRAS | | | | | |
| | Wild type | 182 | 5 | 177 | 0.81 |
| | Mutated | 29 | 0 | 29 | |
| ERBB2 | | | | | |
| | Wild type | 204 | 5 | 199 | 0.40 |
| | Mutated | 7 | 0 | 7 | |

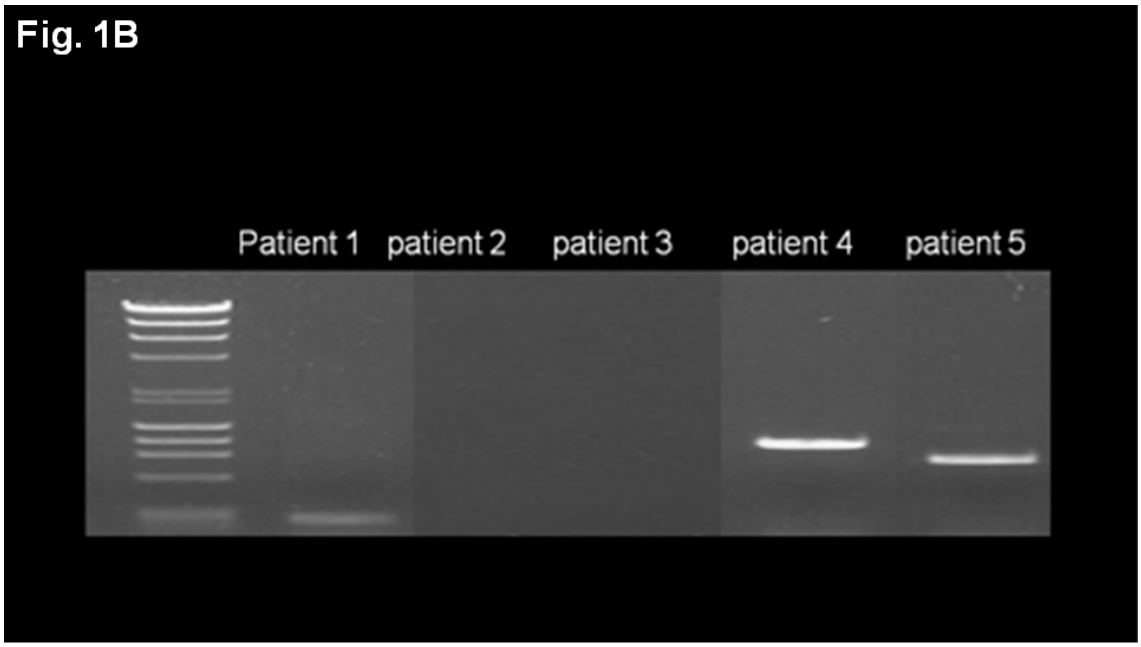
† well plus moderately differentiated adenocarcinomas vs. poorly differentiated adenocarcinomas;
‡ p-I plus II adenocarcinomas vs. p-III plus IV adenocarcinomas

**Table 3. Characteristics of Patients Included in This Study**

| Characteristic | | total (n=313) | Histology | | |
|---|---|---|---|---|---|
| | | | Adenocarcinoma (n=211) | Squamous cell carcinoma (n=75) | Others (n=27) |
| Age (yr; mean) | | | 65.3 | 69.6 | 64.0 |
| Sex | Male | 202 | 111 | 66 | 25 |
| | Female | 111 | 100 | 9 | 2 |
| Smoking | | | | | |
| | Never | 97 | 92 | 3 | 2 |
| | Smoker (ex- / current) | 216 | 119 | 72 | 25 |
| | Pack-year (mean ± SD) | | 28.9 ± 46.6 | 60.1 ± 37.1 | 54.2 ± 33.8 |
| Differentiation | | | | | |
| | Well | 91 | 87 | 3 | 0 |
| | Moderate | 143 | 91 | 52 | 4 |
| | Poor | 80 | 33 | 20 | 21 |
| p-Stage | | | | | |
| | I | 197 | 141 | 43 | 13 |
| | II | 37 | 20 | 13 | 4 |
| | III | 70 | 42 | 19 | 9 |
| | IV | 9 | 8 | 0 | 1 |
| EGFR | | | | | |
| | Wild type | 207 | 106 | 74 | 27 |
| | Mutated | 106 | 105 | 1 | 0 |
| KRAS | | | | | |
| | Wild type | 278 | 182 | 73 | 23 |
| | Mutated | 35 | 29 | 2 | 4 |
| ERBB2 | | | | | |
| | Wild type | 305 | 204 | 74 | 27 |
| | Mutated | 8 | 7 | 1 | 0 |
| EML4-ALK fusion | | | | | |
| | negative | 308 | 206 | 75 | 27 |
| | positive | 5 | 5 | 0 | 0 |

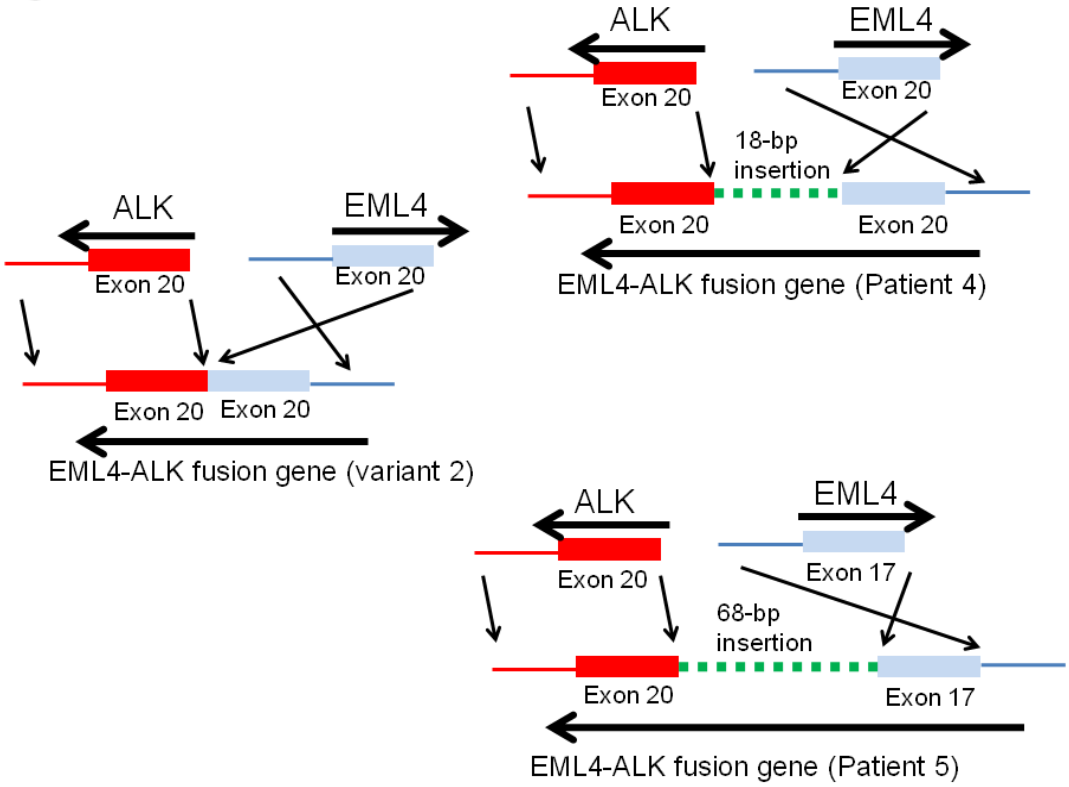



RT-PCR studies using the *EML4* (exon 2)-*ALK* primer set. The λ DNA-*Hin*dIII Digest (BioLabs) and the φX174 *Hae*III digest (Takara Bio) were used as the DNA ladder markers. In each patient (1-5), a single band was amplified, and its size was subsequently determined: 1357-2027 bp band in patient 1, 603-872 bp band in patient 2, 603-872 bp band in patient 3, 2027-2322 bp band in patient 4, and 2027-2322 bp band in patient 5.



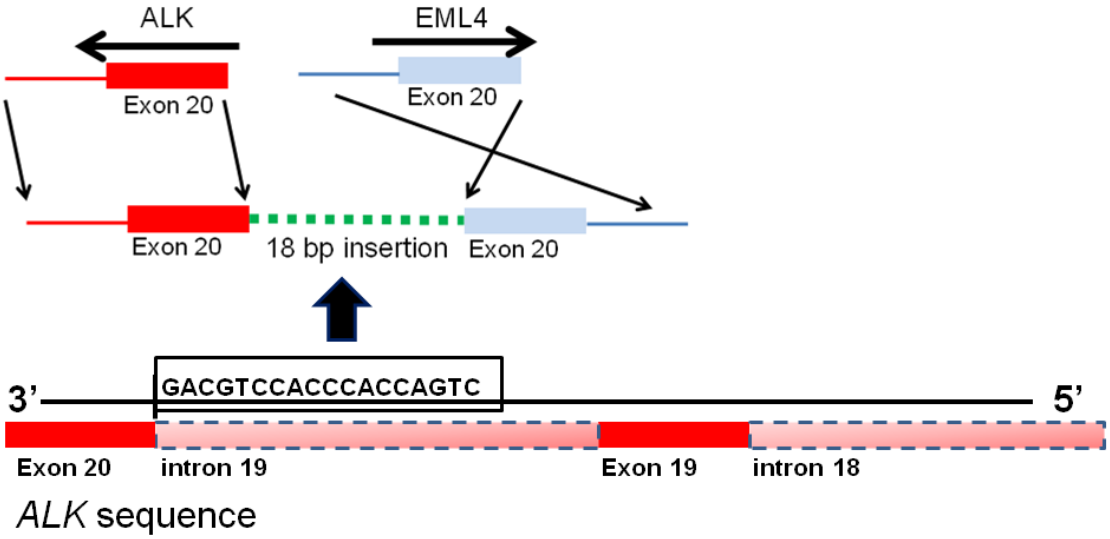

RT-PCR studies using the *EML4* (exon13)-*ALK* primer set. The same DNA ladder markers as above were used. A single band was amplified in three of the five patients, and the sizes were subsequently determined: a <281 bp band in patient 1, 872-1078 bp band in patient 4, and 603-872bp band in patient 5. No RT-PCR products were obtained in patients 2 and 3.

**Fig. 2A**



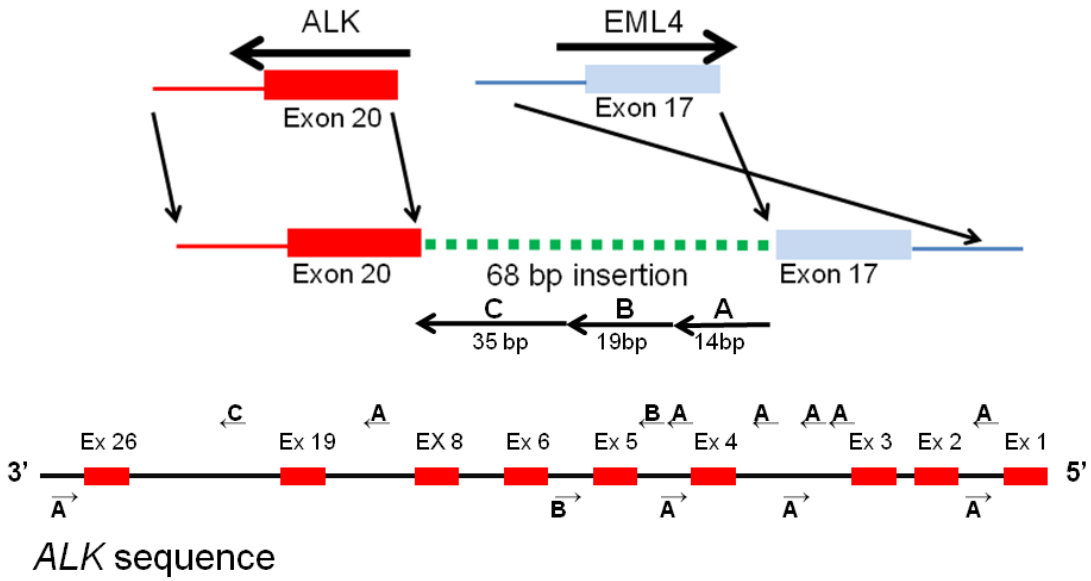

Schema showing two novel *EML4-ALK* fusion variants and variant 2. In patient 4, exon 20 of *EML4* is fused via an 18-bp sequence to exon 20 of *ALK*. This fusion could be classified as a subtype of variant 2. In patient 5, exon 17 of *EML4* is fused via a 68-bp insertion to exon 20 of *ALK*. Compared with variant 2, this fusion type had a full length deletion *EML4* exon 18-20 (275-bp) and a 68-bp insertion. This results in a 207-bp shorter fusion sequence than variant 2, which provided no frame shift.

**Fig. 2B**



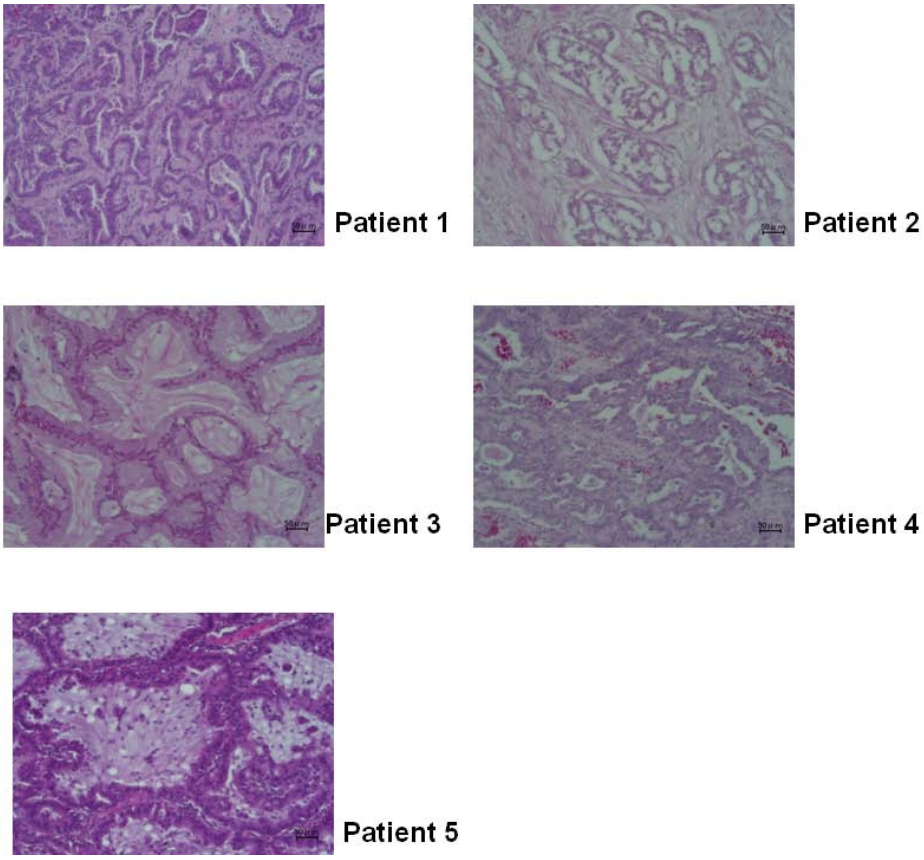

Schema showing the fusion pattern of patient 4 with precise sequence information. The inserted sequence was identical to the last part of ALK intron 19, which indicates that the break point of *ALK* is 18-bp upstream of that in variant 2.

**Fig. 2C**




Schema showing the location of identical sequences composing inserted sequences. Three parts of the inserted sequence are named A, B, and C. Sequences marked as “A” and a left arrow indicate the sequence 5’-AGTCTTGCTCTGTC-3’, and those marked as “A” and a right arrow indicate the reversed sequence 5’-GACAGAGCAAGACT-3’. Sequences marked as “B” with a left arrow indicate the sequence 5’-TCCCAGGCTGGAGTGCAGT-3’, and those marked by “B” with a right arrow indicate the reversed sequence. Sequence marked as “C” with a left arrow indicates the 35-bp sequence 5’-GGCAATTTACACATTTCAATTCATTCGATCCTCAG-3’.

**Fig. 3**




Histopathological studies of *EML4-ALK* fusion-harboring tumors. All of the carcinomas were adenocarcinomas. One was an acinar adenocarcinoma (patient 2), and the other four were classified as adenocarcinomas of mixed subtypes. In patients 3 and 4, acinar characteristics were evaluated to be predominant. The acinar adenocarcinoma in patient 2 was classified as moderately differentiated, and the others were all classified as poorly differentiated (grade 3) types.